

A DNA Polymerase III Holoenzyme-like Subassembly from an Extreme Thermophilic Eubacterium

Charles S. McHenry^{1*}, Mark Seville² and Millard G. Cull²

¹Department of Biochemistry and Molecular Genetics and the Molecular Biology Program University of Colorado Health Sciences Center, Denver CO 80262, USA

²Enzyco, Inc., 191 University Blvd., Ste. 305, Denver CO 80206, USA

We have purified a novel DNA polymerase from *Thermus thermophilus*. This was enabled by use of general gap filling assays to monitor polymerase activity and cross-reactive monoclonal antibodies against the α catalytic subunit of *E. coli* DNA polymerase III holoenzyme to distinguish a novel polymerase from the well characterized DNA polymerase I-like *Thermus thermophilus* DNA polymerase. Two proteins migrating with the polymerase after three chromatographic steps were isolated and subjected to partial amino acid sequencing. The amino termini of both were homologous to the two products of the *E. coli* *dnaX* gene, the γ and τ subunits of the DNA polymerase III holoenzyme. Using this information and sequences conserved among *dnaX*-like genes, we isolated a gene fragment by PCR and used it as a probe to isolate the full length *Thermus thermophilus* *dnaX* gene. The deduced amino acid sequence is highly homologous to the DnaX proteins of other bacteria. Examination of the sequence permitted identification of a frameshift site similar to the one used in *E. coli* to direct the synthesis of the shorter γ DnaX-gene product. Based on this information, we conclude that a conventional replicase exists in extreme thermophilic eubacteria. The general biological and practical technological implications of this finding are discussed.

© 1997 Academic Press Limited

Keywords: thermophile; *T. thermophilus*; DNA polymerase; holoenzyme; DNA replication

*Corresponding author

Introduction

Escherichia coli, an often studied prototype for general eubacterial mechanisms, encodes three distinct DNA polymerases. DNA polymerase I is the most abundant polymerase and is responsible for some types of DNA repair, including a repair-like reaction that permits the joining of Okazaki fragments during DNA replication. DNA polymerase III comprises the catalytic core of the *E. coli* replisome. There are approximately 400 copies of DNA polymerase I per cell, but only 10 to 20 copies of Pol III (Kornberg & Baker, 1992; Wu *et al.*, 1984). The low abundance of Pol III and its relatively feeble activity on gapped DNA templates typically used as a general replication assay delayed its discovery until the availability of mutants defective in DNA polymerase I (Kornberg & Geftter, 1972).

The catalytic subunit of Pol III is distinguished from other polymerases as a component of *E. coli* major replicative complex, apparently not by its

intrinsic catalytic activity, but by its ability to interact with other replication proteins at the fork. These interactions confer upon the enzyme enormous processivity. Once the DNA polymerase III holoenzyme associates with primed DNA, it does not dissociate for over 40 minutes, the time required for the synthesis of the entire 4 Mb *E. coli* chromosome (McHenry, 1988). Studies in coupled rolling circle models of the replication fork suggest the enzyme can synthesize DNA 150 kb or longer without dissociation *in vitro* (Mok & Marians, 1987; Wu *et al.*, 1992). The essential interaction required for this high processivity is an interaction between the α catalytic subunit and a dimer of β , a sliding clamp processivity factor that encircles the DNA template like a bracelet, permitting it to rapidly slide along with the associated polymerase, but preventing it from falling off (LaDuca *et al.*, 1986; Kong *et al.*, 1992).

The β_2 bracelet cannot spontaneously associate with high molecular mass DNA, it requires a multiprotein DnaX-complex to open and close it around DNA using the energy of ATP hydrolysis (Wickner, 1976; Naktinis *et al.*, 1995; Dallmann

Abbreviations used: Pol, polymerase; Fr, fraction.

et al., 1995). In *E. coli*, the *dnaX* gene encodes two proteins, τ and γ . γ is generated by a programmed ribosomal frameshifting mechanism five-sevenths of the way through *dnaX* mRNA, placing the ribosome in a -1 reading frame where it immediately encounters a stop codon (Flower & McHenry, 1990; Blinkowa & Walker, 1990; Tsuchihashi & Kornberg, 1990). In *E. coli*, the DnaX-complex has the stoichiometry $\gamma_2\tau_2\delta_1\delta'_1\chi_1\psi_1$ (Dallmann & McHenry, 1995). The τ protein contains an additional carboxyl-terminal domain that interacts tightly with the polymerase, holding two polymerases together in one complex that can coordinately replicate the leading and lagging strand of the replication fork simultaneously (McHenry, 1982, 1988; Studwell-Vaughan & O'Donnell, 1991).

Pol IIIs are apparently conserved throughout mesophilic eubacteria. In addition to *E. coli* and related proteobacteria, the enzyme has been purified from the firmicute *Bacillus subtilis* (Low *et al.*, 1976; Hammond & Brown, 1992). By inference from DNA sequence information available from the proliferation of bacterial genomes sequenced, Pol III exists in organisms as widely divergent as *Caulobacter*, *Mycobacteria*, *Mycoplasma*, *B. subtilis* and *Synechocystis*. The existence of *dnaX* and *dnaN* (structural gene for β) is also apparent in these organisms. These general replication mechanisms are conserved even more broadly in biology. Eukaryotes, although they do not contain polymerases homologous to Pol III, contain special polymerases devoted to chromosomal replication, β -like processivity factors (PCNA) and DnaX-like ATPases (RFC, Activator I) that assemble processivity factors on DNA (Yoder & Burgers, 1991; Brush *et al.*, 1995; Uhlmann *et al.*, 1996).

In spite of the near ubiquity of Pol IIIs and their associated factors required to function as a replicase, no enzymes of this class have been found in thermophilic eubacteria, especially in the closely related *Thermus aquaticus* and *Thermus thermophilus* that has been carefully studied in several laboratories. This raised a paradox. Is the mechanism of replication somehow different in thermophiles? Here, we describe studies where we definitively establish the existence of a Pol III holoenzyme-like

subassembly that contains, minimally, a Pol III α subunit and associated τ and γ subunits. The latter is apparently produced by a frameshifting mechanism similar to that used in *E. coli*. The technical implications of the existence of a highly processive replicase in an extreme thermophile are also discussed.

Results

Detection and partial purification of *T. thermophilus* Pol III and associated proteins

A survey of 12 monoclonal antibodies directed against the 130 kDa α subunit of the *E. coli* Pol III holoenzyme revealed a subset that reacted with a protein of approximately the same size in Western blots of *T. thermophilus* extracts. These antibodies were used as a tool to distinguish Pol III-like polymerases from the characterized *T. thermophilus* polymerase during protein fractionation procedures.

Having established that a roughly 130 kDa protein cross-reacted with anti-*E. coli* α monoclonals, we optimized a lysis procedure and ammonium sulfate fractionation that provided α . We found that use of a *T. thermophilus* strain carrying an S-layer *T. thermophilus* mutation (*slpA*) made the cells suitably sensitive to lysozyme that a modification of the standard *E. coli* gentle lysis procedure (Cull & McHenry, 1995) provided a satisfactory method to produce Fraction I (Table 1). We optimized ammonium sulfate precipitations to provide a nearly quantitative precipitation of the candidate *T. thermophilus* α subunit (judged by Western blots) while removing as much contaminating protein as possible (Fr II; Table 1). Fraction II was dissolved and subjected to BioRex70 cation exchange chromatography. This method did not result in resolution of DNA polymerases, but did remove significant levels of contaminating protein that flowed through the column. A 25-fold increase in gap-filling polymerase specific activity was achieved (Fr III, Table 1).

Hydrophobic interaction chromatography was used to resolve a novel minor polymerase peak

Table 1. Purification of *T. thermophilus* Pol III and associated proteins

Fraction	Units ($\times 10^3$)	Protein (mg)	Specific activity (units/mg)
I. Lysis	ND*	73,600	ND
II. Ammonium sulfate	2200	6540	360
III. BioRex-70	6700	686	10,000
IV. ToyaPearl-Ether	1400	54	25,000
V. Q-Sepharose	760*	4.8	160,000*

* Fraction I was not assayed due to a non-linear and unreliable response, presumably due to nuclease or inhibitory contaminants.

* Assays were conducted at 37°C. The specific activity is fivefold higher (800,000 units/mg) if assayed at 60°C.

* The reported yield is normalized to a complete preparation using all of the material from Fr IV. We used only 50% of Fr. IV in the preparation of Fr. V.

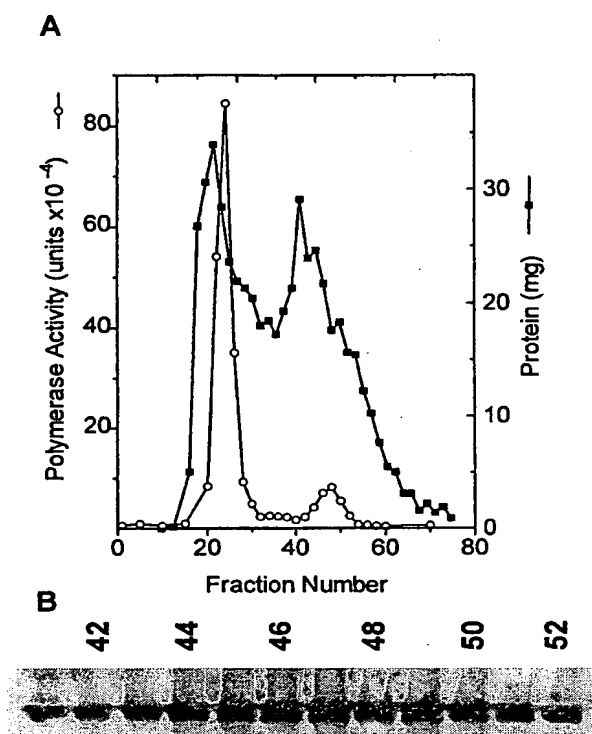


Figure 1. Separation of the major *T. thermophilus* DNA polymerase from Pol III. A, Hydrophobic interaction chromatography (ToyaPearl-Ether 650 M) of Fr III conducted and analyzed as described in Methods. B, Fractions from a second minor peak of activity were subjected to the Western blotting procedure as described in Methods to reveal a band that comigrated with *E. coli* Pol III α subunit. Note that the intensity of the band by Western blots parallels polymerase activity in the second peak.

from the majority of *T. thermophilus* DNA polymerase activity. Fr III was dissolved and precipitated onto ToyaPearl-Ether 650 M beads by adjusting the ammonium sulfate concentration to 1.8 M. Protein was eluted with a decreasing ammonium sulfate gradient (Figure 1A). The major *T. thermophilus* DNA polymerase eluted early in the gradient and did not react with anti- α monoclonal antibodies. Peak 2, a minor peak comprising 9% of the polymerase activity, bound more tightly to the column (Fr IV, Table 1). Fractions subjected to a Western blotting procedure revealed that the cross-reactive 130 kDa band eluted with an intensity parallel to the polymerase activity in peak 2 (Figure 1B). On this basis, we identify the second peak of activity as a Pol III-like polymerase, distinct from the characterized *T. thermophilus* DNA polymerase.

To provide a higher level of purification to enable direct examination of the components, we

developed an additional chromatographic step. Fraction IV was subjected to chromatography on Q-Sepharose to yield a single peak of polymerase activity and a sixfold increase in specific activity (Figure 2A; Fr V, Table 1). An SDS/polyacrylamide gel of the peak fractions showed that a 130 kDa protein eluted in parallel to the activity profile (Figure 2B). A 43 kDa protein eluted later and did not parallel activity and is presumably a contaminant. Two additional proteins of $63(\pm 6)$ and $50(\pm 5)$ kDa eluted roughly in parallel with the 130 kDa protein, providing candidates for

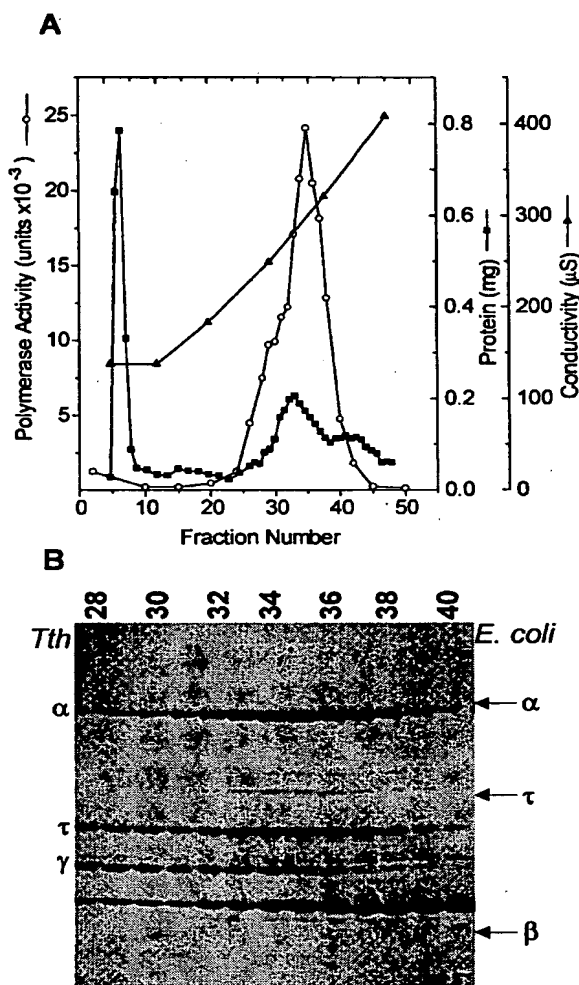


Figure 2. Purification of *T. thermophilus* Pol III by anion exchange chromatography on Q-Sepharose. A, Chromatography was conducted and samples analyzed as described in Methods. B, Samples (60 μ l) were subjected to SDS/polyacrylamide gel electrophoresis (10% gel) as described. The identities of the purified *T. thermophilus* subunits are designated in the left margin; the migration positions of *E. coli* DNA polymerase III holoenzyme markers are indicated in the right margin.

E. coli	3	LYRRFRPLTFQEVVGQEH	20
		L R++RP TF +VVGQEH	
Tth gamma	1	SALYRRFRPLTFQEVVGQE	19
Tth tau	1	SALYRRFRPLTFQEVVGQEH	20
		ALYR FRP F++VVGQEH	
B. subtilis	5	ALYRVFRPQRFEDVVGQEH	23

Figure 3. Amino-terminal sequence of the isolated candidate γ and τ *T. thermophilus* *dnaX* gene products and comparison to the homologous sequences of *E. coli* and *B. subtilis*. (*T. thermophilus* abbreviated as *Tth*). Letters are placed in intervening lines where identical matches occur. A (+) indicates a similarity in aligned residues.

polymerase III-associated proteins, possible subunits of a *T. thermophilus* Pol III holoenzyme.

Identification of *T. thermophilus* homologs of the *E. coli* Pol III holoenzyme γ and τ subunits

The 63 and 50 kDa proteins associated with the *T. thermophilus* Pol III α subunit were separated by SDS PAGE, blotted to a membrane and subjected to amino-terminal sequencing. The 19-residue sequence obtained from the 50 kDa protein was identical to the first 19 residues of the sequence obtained for the 63 kDa protein, consistent with both being the product of the same gene (Figure 3). The sequenced regions were 60% identical to *E. coli* DnaX, the only bacterial DnaX protein that has been directly characterized on the basis of activity. It was 70% identical to *B. subtilis* DnaX. This is highly suggestive that the 63 and 50 kDa proteins are the *T. thermophilus* γ and τ subunits of Pol III holoenzyme. We also obtained an internal peptide sequence from the isolated 63 kDa band. The sequence: ARLPLAQAHFGVEEVVLVEGE did not show any detectable homology to known *dnaX* genes.

Cloning and sequencing of the *T. thermophilus* *dnaX* gene

To definitively test our hypothesis that the 63 and 50 kDa proteins are products of the *T. thermophilus* *dnaX* gene, we used a PCR approach to obtain a fragment of the gene. The gene fragment was used as a probe to obtain the full length of *T. thermophilus* DnaX, enabling its full sequence to be determined. From the N-terminal sequence, a peptide was selected, FQEVVGQ, that would provide a PCR primer with the least degeneracy with all possible codons represented (512-fold). To provide an internal primer, we exploited a KTLEEP amino acid sequence common to *E. coli*, *B. subtilis* (Carter *et al.*, 1993; O'Donnell *et al.*, 1993) and many other eubacteria. From this consensus sequence a 17-mer primer with 256-fold degeneracy was designed that included all possible combinations of codons. Use of these primers in PCR reactions yielded a

product of 388 nucleotides, close to the spacing of the corresponding regions in the *E. coli* *dnaX* gene (393 nucleotides). The PCR product was cloned and sequenced. The deduced amino acid sequence was 42% identical to the corresponding segment of the *B. subtilis* *dnaX* gene.

Since we had a large probe that was identical to the gene, we decided to use a directed approach toward cloning the full length *T. thermophilus* *dnaX* rather than screening of a full library. We performed a Southern blot against digested *T. thermophilus* DNA and found that the PCR probe hybridized to a 7.1 kb *Pst*I fragment that was large enough to contain the entire *T. thermophilus* *dnaX* gene. *Pst*I-digested DNA was subjected to electrophoretic separation and the region corresponding to the sought 7.1 kb fragment was extracted and cloned. The resulting colonies were screened by colony hybridization and a candidate selected. It was sequenced, resulting in detection of a 2142 bp open reading frame (Figure 4A). Within the candidate open reading frame GUG (the 186th codon) was identified as the actual initiation codon from the sequence of the amino terminus; it was immediately followed by the experimentally determined amino-terminal sequence reported in Figure 3. GUG is frequently used as an initiation codon in *T. thermophilus*. Methionine aminopeptidases would be expected to cleave off the terminal methionine (Sherman *et al.*, 1985), revealing serine as determined experimentally. The translation of the deduced sequence of the *T. thermophilus* τ subunit is shown in Figure 4B. We observe a perfect match with the internal peptide sequence obtained from sequencing of the 63 kDa candidate γ subunit between residues 428 and 450, further confirming that we isolated the structural gene for the protein associated with *T. thermophilus* Pol III. Alignment of the resulting sequence with the identified *E. coli* *dnaX* gene and putative homologous genes from other eubacteria reveal a high level of identity in the amino-terminal region of all bacteria, confirming the identity of the isolated gene to be a *dnaX* homolog (Figure 4C). The molecular mass of the predicted *T. thermophilus* τ subunit is 58 kDa, in reasonable agreement with the 63(\pm 6) kDa determined from SDS PAGE.

In *E. coli*, the shorter γ product of the *dnaX* gene is produced by a translational frameshifting mechanism at the sequence A AAA AAG containing two adjacent lysine codons read by the Lys UUU anticodon tRNA. A potential frameshift site that exploits adjacent AAA Lys codons flanked by A residues (A AAA AAA A) is observed at codons 451 and 452 (Figure 4A). Frameshifting into the -1 reading frame at this site would result in a 51 kDa protein, close to the 50 kDa candidate observed in Fr. V (Figure 4D). If, instead, a +1 frameshift occurs, a 50 kDa protein would be produced, indistinguishable within experimental error from the -1 frameshift alternative.

A

TCCGCCAGGCGCTTCCCGTGAGGTAGAGGGCCATCTCCACCACCTCGAGCTGCTCC
 GGGGGCGTCACCCGAAGCCCCCCCCACGAAGCGGCTCTCGTACCCAGGCGCCCAAGC
 CAAGCCCCGATCTCGGGCCCCCCCCCGTGACGAGAACCAAGGGACCAAGGCTAGGC
 GGCAAGCTCGTCCAATAGGGCTCCGCCCCCTAAGGCTTCTCCACCTTCACCAAAA
 GGGCTCACTCAAGGTAGACCACCTCTCGGGCAGGCCCTCTCTCTCGGGCTCCAG
 GAGGTGCGCGAAGCCAGAAGGGCGTGGTCCGGGTGGAGGCCGAAGCGGGCGGCGAG
 GGCCCGGACCGGTGCTCGGGCGGGAGGGCTGGGCCCTCGTCAGGAGGGGGCGGAG
 GTCGGCGGGGGCTTCAGGGCCGTCCCAAGCTCCGGACCCGCGCGAAGCTCCGCCCTC
 ACCTCCCCCCCCCTCGGCCAGGAGGAAGAGGAGGTCTCTCTCGCTTAGTACCAGGAAGGC
 CAAGGCTGCCCCAGGCGGGAAGCTCCCGCAGGAAGGCCCGGATGGCCTCCGGGTCC
 TCTCCGCTCGAGGGCTCTGTGTAGAGGCTCGTCCAAGGGGGGTGCGGGACCAAGT
 AGACCCCGTCCCTCCCGTGGCGCCCCCTCGGCCAGGCCGCCACCGCTCCAGGGGGGC
 CTGACGTGCAAGGAGAGGAAGCTCCGCACCAACCGCTATAGTACCCCTTCTGAGCGCC
 CTCTACCGCCGCTTCCGCCCTCACCTTCCAGGAGGTGGTGGGCGAGGACACGTG
 AAGGAGCCCTCTCTCAAGGCCATCCGGAGGGGAGGCTCGCCAGGCCCTACCTCTC
 TCCGGGCCAGGGGCGTGGCAAGACCACACGGCGAGGCTCTCGCCATGGCGGT
 GGGGTGCCAGGGGAAGACCCCTTGGGGGTCTGCCCCACTGCCAGGCGGTGCA
 GAGGGCGCCCAACCGGACGTGGTGGAGATTGACGCGCCAGCAACAACCTCGGTG
 AGGACGTGCGGGAGCTGAGGGAAGGATCCACCTCGCCCCCTTTCTGCCCCAGGA
 AGGTCTTCATCTGGACGAGGCCACATGCTCTCCAAAAGCGCTTCAACGCCCTCT
 CAAGACCTGGAGGAGCCCCGCCCCACGTCTCTCTGCTTTCGCCACCAACCGAGCC
 CGAGAGGATGCCCCCACCATCTCTCCCGCACCAGCACTTCCGCTTCCGCCGCTC
 ACGGAGGAGGAGATCGCCTTAAGCTCCGCGCATCTGGAGGCCGTGGGGCGGGA
 GGGGAGGAGGAGGCCCTCTCTCTCGCCCGCTGGCGGACGGGGCCCTTAGGGA
 CGCGGAAAGCCTCTGGAGCGCTTCTCTCTCGGAAGGCCCTTCAACCGGAAGGA
 GGTGGAGCGCGCCTAGGCCTCCCCCAGGGAGGCCCTGGCCGAGATCGCGCCTC
 CCTCGCGAGGGGAAAACGGCGGAGGCCCTGGGCCTCGCCCGCGCCTCTACGGGG
 AAGGTTACGCCCCGAGGAGCCTGGTCTCGGGCTTTTGGAGGTGTTCCGGGAAGGCC
 TCTACGCCGCTTTCGGCTCGCGGGAACCCCTTCCCGCCCCGCCCCAGGCCCTGAT
 CGCCGCCATGACCGCCCTGGACGAGGCCATGGAGCGCTCGCCCGCCGCTCCGACGC
 CTTAAGCCTGGAGGTGGCCTCTGGAGGCGGGAAGGGCTTGGCCGCGAGGCCCT
 GCCCCAGGCCACGGGCGCTCCCCCCCCAGAGGTGGGCCCAAGCCGGAAGCCCCC
 GGCCCCGAACCCCAAGGCCCGAGGAGGCCCGACCTGCGGGAGCGGTGGCGGG
 CCTTCTCGAGGCCCTCAGGCCACCTACGGGCCTTCTGCGGGAGGCCCGCCCG
 AGGTCCGGGAAGGCCAGCTCTGCCTCGCTTCCCCGAGGACAAGGCCCTTCACTACC
 GCAAGGCCTCGGAACAGAAGGCGAGGCTCTCCCCCTGGCCAGGCCCATTCGGGG
 TGGAGGAGGTCTCTCTGCTTGGAGGGAGAAAAAAGCCTGAGCCCAAGGCC
 CGCTCGGCCACCTCTGAAGCGCCCGACCCCGGGCCTCCGAGGAGGAGGT
 GAGGCGGAGGAAGCGCGGAGGAGGCCCGGAGGAGGCCTTGAAGCGGGTGGTCC
 GCCTCTGGGGGGCGGGTCTCTGGGTGCGCGGCCAGGACCCGGGAGGCGCG
 GAGGAGGAACCCCTGAGCCAAGACGAGATAGGGGGTACTGGTATATAAGGGGGA
 TGACGCGGACACCGACCTCGACAAAGAGACCGTGGACAACATCTCAAGCGCTCC
 GCCGTATTGAGGGCCAGGTGCGGGGCTCCAAAGATGGTGGCCGAGGGCCGCC
 TGCGACGAGGTCTCACCCAGATGACCGCCACCAAGAAGGCCATGGAGGCGCGGC
 CACCCTGATCTCCACGAGTCTCTGAACGTCTGCGCGCCGAGGTCTCCGAGGGCAA
 GGTGAACCCCAAGAAGCCCGAGGAGATCGCCACCATGCTGAAGAAGTTCATCTAGAT
 GGGTCCGCTTCCGGGGCGCCTCCGGCGCCTCTCCGGGCCCTTCTCGCCAGGAGGC

B

SALYRRFRPLTFQEVVGOEHVKEPLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMA
 VGCQGEDPPCGVCPHCQAVQRGAHPDVVEIDAASNNVEDVRELREIRHLAPLSAPRKV
 FILDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPPTILSRQHFRFRRLTEEEIA
 FKLRRILEAVGREAEELALLLARLDGALRDAESLLERFLLEGLPLTRKEVERALGLPPR
 EALAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLELVFREGLYAAFLAGTPLPA
 PPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPPPEVGPKE
 SPPAPEPPRPEEAPDLRERWRAFLAALRPTLRAFVREARPEVREGQLCLAFPEDKAFHYR
 KASEQKARLLPLAQAHFGVEEVVLVLEGEKKSLSRPRSAPPPEAPAPPGPPEEEVEAE
 EAAEEAPEEALRRVVRLLGGRVLWVRPRPTREAPPEEPLSQDEIGGTGI

Figure 4(A–B) (legend opposite)

C

Tth	1SALY	RRRRLTFQE	VVGQEHVKEP	LLKATREGRL	ACAYLFSGPR	GVGKT
E.coli	1	...MSYQVLA	RKWRPQTFAD	VVGQEHVLT	LANGLSLGR	HHAYLFSGTR	GVGKT
B.sub	1	...MSYQALY	RVERPQRFED	VVGQEHITKT	LQNALQKKF	SHAYLFSGPR	GVGKT
Mycopl	1	...MRKVLY	OKYRPTKFS	TVGQDSIKRI	IVNATQDQ	PHGYTFAGER	GVGKT
Caulo	1	EKGDAITVLA	RKYRPTKFS	LIGQBAVVRT	LANAFSTGRI	AHAFMLTCVR	GVGKT
Syn.sp.	1	...MANEPLH	HKYRPTKFS	LVGQTAHAAT	LSNAIFQERI	VPAYLFTGPR	GVGKT
consensus		y L	rkfRP tF d	vvgQe v	l nai	rl haylfsG R G	GKT

Tth	50	TTARL	LAMAVCGQGEDPPC	GVCPHCQAVQ	RGAPDVVEI	DAASNNSVED
E.coli	53	SIARL	LAKGNCETGITATPC	GVCDNCRELE	QGRFVDLIEI	DAASRTKVED
B.sub	53	SAAKI	FAKAVNCEHAPVDEPC	NECAACKGIT	NGSISDVIEI	DAASNNGVDE
Mycopl	52	TFARI	LAKAINCLNWNGDVC	NOEACQAIN	SNSALDVIEI	DAASKNGIND
Caulo	56	TTARL	LARALNYETD	TVKGPSVDLT	TEGYHCRSTI	EGRHMDVIEI	DAASRTKVED
Syn.sp.	53	SSARI	LAKSLNCIAG	DR..PTATPC	GCATCRAIT	NGSALDVIEI	DAASNTGVIN
consensus		t Arl	lakavnc	pc	c Crai	g vDvveI	DAAS ved

Tth	99	VRELREIRHL	AFLSAPRKVF	ILDEAHMLSK	SAFN	ALLKT	LEEP	PFHV	LFVF
E.coli	104	TRDLLDNVOY	APARGREKVY	IIDEVHMLSR	HSFN	ALLKT	LEEP	PEHV	KFIL
B.sub	104	IRDLRDKVKF	APSAYTKVY	IIDEVHMLSI	GAFN	ALLKT	LEEP	PEHC	IFIL
Mycopl	102	IRELAENVFN	LPFKFKKKVY	ILDEAHMLTP	QSFS	ALLKT	LEEA	PDYV	EFIL
Caulo	111	IRELLDGVRY	APVEAREKVY	IIDEVHMLST	AAFN	ALLKT	LEEP	PFHA	KFIL
Syn.sp.	106	IRELIERAQF	AEVQCQYKVY	VIDECLTGDS	QVLTRNGLMS	IDNEQIKGEE	VLST		
consensus		vRel e v y aP	rFKVY	iide hmls	fn allkt	leep	p h	lfvf	

Tth	150	ATTEPE	RNPPTILSRT	QHFR....FK	RTBEEIAFK	IRRIIEAVGR	EAEAEALL
E.coli	155	ATTDPO	KIPVTILSRC	LQFH....LK	AEDVEQIRHQ	HEHINEEHI	AHEPRALCL
B.sub	155	ATTEPH	KIPVTILSRC	ORFD....FK	RITSQAIVGR	MNKIVDAEQ	QVEEGSEI
Mycopl	153	ATTEFN	KIPVTILSRC	QSFF....FK	CHTNDLIQOR	DAEVAAKESI	KITTDALVK
Caulo	162	ATTEIR	KIPVTILSRC	ORFD....LR	RVEPDVLVKF	FDRISAKEGA	RIEMDALAI
Syn.sp.	160	NETLQQ	WEYKQVLRWL	DGEKQTLST	KTKNSTVRC	ANHLIRTEOG	WTRAEITP
consensus		atTe	kmpvtilsrc	qhF	r rl e i k l ril e	e eal l	

Tth	201	L APLADGALRD	AESLLERFIL	L...EGP	LIR	KEVERALCLP	PE.....
E.coli	206	L ASAAEGSLRD	ALSLLDQATA	SG..DGO	ST	QAVSAMIGTL	DDDQALSLVE
B.sub	206	L ASAADGCMRD	ALSLLDQALS	FS..GDIE	KV	EDALLITGAV	SQ.....
Mycopl	204	L ADLAAGSLRD	GLSLLDQISN	FSE.SKIT	SL	ADVEKTFNLL	DREQ.....
Caulo	213	L ASAAEGSVRD	GLSLLDQATV	QTERGQTVIS	TVVRDMLGLA	DR.....	
Syn.sp.	215	G MUILSPASVD	VDNLSQSTAL	TAS.LGGL	SG	AINYEAINTD	KNTTSLSL
consensus		l ar adg lrd	alsLldq l		it v lg	r	

D

GAAAAAAAAAGCCTGAGCCCAAGCCCCGCTCGGCCCCACCTCCTGA
 E K K S L S P R P R S A P P P
 -1 E K K K P E P K A P L G P T S
 +1 E K K A

Figure 4. Sequence and analysis of the full length *T. thermophilus* *dnaX* gene. A, DNA sequence of *T. thermophilus* *dnaX* and flanking sequences. Note that the initiating GTG was the 186th codon in a long open reading frame and could only be conclusively identified by use of the amino-terminal sequence of protein purified from *T. thermophilus*. The preceding sequences in the open reading frame are italicized. Also in bold and underlined is the AAAAAA sequence that, by analogy to *E. coli*, is probably the frameshifting site that permits synthesis of the shorter γ product. The potential stop codons for γ in the +1 (first) and -1 reading frame are dotted/underlined. The stop codon for the full length τ translation product is underlined. B, The translation of the τ gene product of *T. thermophilus* *dnaX*. The protein sequences directly determined from the isolated *T. thermophilus* DnaX-proteins are underlined. C, Comparison of *T. thermophilus* *dnaX* with homologous eubacterial *dnaX* sequences. To provide the most useful and concise presentation, only examples from widely divergent organisms are presented. Organisms listed from the top are: *T. thermophilus*; Chloroflexaceae/Deinococcaceae group), *E. coli* (proteobacteria group, gamma division), *B. sub.* (*B. subtilis*, firmicute group, low G+C Gram-positive bacteria division), *Mycopl* (*Mycoplasma pneumoniae*, firmicute group, mycoplasma division), *Caulo* (*Caulobacter crescentus*, proteobacteria group, alpha division), *Syn.sp.* (*Synechocystis* sp., Cyanobacteria group, (blue-green algae)). Sequences shaded in black are identical among the indicated bacteria; sequences shaded in gray are similar. The alignment was not shown beyond *T. thermophilus* DnaX residue 240 because it was insignificant. D, Deduced sequence of *T. thermophilus* γ subunit depending on whether the frameshift is -1 (as in *E. coli*) or +1 as would be permitted from the sequence as described in Discussion. The sequence preceding the sequence shown should be identical for τ and γ .

Discussion

Evidence from biochemistry and genomics suggests that the mechanism of DNA replication is uniform throughout biology. Yet, studies in thermophilic eubacteria have not led to the discovery of Pol III holoenzyme-like replicases. Armed with a bank of specific monoclonal antibodies against the α subunit of Pol III and our present knowledge of the properties of Pol III complexes, we took a directed approach in a search for Pol III holoenzyme-like complexes in *T. thermophilus*.

The initial cation exchange step was used because polymerases, in general, tend to interact strongly with cation exchangers, presumably through active site nucleic acid binding residues since the overall proteins are acidic. This resulted in isolation of an enriched polymerase fraction, but did not resolve the Pol III antibody reactive fraction from the majority of polymerase activity. *E. coli* Pol III holoenzyme binds tightly to hydrophobic columns, affording an effective purification step (McHenry & Kornberg, 1977). Thus, we attempted hydrophobic interaction chromatography in the hope of resolving *T. thermophilus* Pol III from the smaller and, presumably, more hydrophilic DNA polymerase I-like activity. The chromatographic profile, based on gap-filling polymerase activity, revealed a major loosely bound peak and a more tightly bound minor peak. The second tightly bound polymerase cross-reacted specifically with *E. coli* anti- α monoclonals; the first peak did not. An additional chromatographic step on the resolved polymerase afforded material where the candidate α and two apparently comigrating components of 63 and 50 kDa comprised approximately 50% of the protein resolved on an SDS gel. The 63 kDa protein chromatographed in parallel with the polymerase. The 50 kDa protein was more strongly represented toward the early portion of the peak.

The associated 63 and 50 kDa proteins were subjected to amino-terminal sequencing, revealing identical termini, consistent with their being the product of the same gene. The 20-residue terminal sequence was 60 and 70% identical to amino-terminal sequences of the *E. coli* and *B. subtilis* DnaX proteins, respectively. The association of a *T. thermophilus* DnaX homolog with a protein that cross-reacts with anti-pol III α monoclonal antibodies and is the same molecular mass as the *E. coli* Pol III α subunit in a 50% pure protein preparation, we argue, establishes the existence of a Pol III holoenzyme-like complex in *T. thermophilus*. Although we do not have a holoenzyme specific assay that enables rigorous quantitation of the extent of purification, Pol III of *E. coli* needed to be purified 30,000-fold (McHenry & Crow, 1979). If the *T. thermophilus* polymerase III has been purified 15,000-fold, the chances of a *dnaX* homolog co-purifying by chance is exceedingly remote. Since we do not yet possess the *T. thermophilus* β subunit, we can-

not determine whether the isolated complex is active in replication complex assembly.

To confirm the identity of the associated DnaX-like components, we isolated their structural gene by a reverse genetic approach. We isolate a PCR fragment of the gene using the amino-terminal sequence experimentally obtained and an internal peptide sequence that is conserved among eubacterial DnaX proteins. We obtained a PCR fragment of the predicted size that was highly homologous to *dnaX* genes in the fragment between the regions corresponding to the PCR primers, eliminating the bias imposed by primer selection. The PCR-generated probe led to the isolation of a full length gene (Figure 4A) that is highly homologous to eubacterial *dnaX* genes. Every region conserved among other eubacterial *dnaX* genes is found represented in *T. thermophilus* *dnaX*. These sequences include the Walker-type ATP binding site, represented by GVGKTTT in *T. thermophilus* *dnaX*, and highly conserved EIDAAS and FNALLKTLLEP sequences (Figure 4C). The conservation beyond the first 220 residues falls off for all *dnaX* genes. Thus, the internal peptide sequence obtained that starts at residue 427 (second underlined sequence in Figure 4B) was useful in confirming that the DNA sequence was in the correct reading frame and providing further confirmation that we have isolated the structural gene for the same proteins isolated in association with *T. thermophilus* Pol III.

The specific activity of the highly purified *T. thermophilus* Pol III was 800,000 units/mg at 60°C. The specific activity of pure *E. coli* Pol III is 2.5×10^6 at 30°C (Kim & McHenry, 1996). After correction for the presence of contaminants and the associated τ and γ proteins, the specific activity of *T. thermophilus* Pol III core is ca 2×10^6 , close to the activity of its *E. coli* counterpart when compared at temperatures slightly below their optimal growth temperature.

Our studies also provide information regarding a conservation of a frameshifting mechanism to generate related ATPases that, by analogy to *E. coli*, can both assemble a β processivity factor onto primed DNA. Both a 63 kDa τ subunit that has a molecular mass consistent with its being a full length *dnaX* translation product and a 50 kDa γ subunit that likely arises by translational frameshifting were detected in enzyme purified from *T. thermophilus* extracts. Examination of the *dnaX* DNA sequence is consistent with this hypothesis. In *E. coli*, ribosomes frameshift at the sequence A AAA AAG into a -1 frame where the lysine UUU anticodon tRNA can base-pair with six As before elongating (Flower & McHenry, 1990; Blinkowa & Walker, 1990; Tsuchihashi & Kornberg, 1990). In *T. thermophilus*, the putative frameshift site has the sequence A AAA AAA A, which would enable either a +1 or -1 frameshift. The +1 frameshift product would extend only one residue beyond the Lys-Lys encoding sequence where the frameshift occurs, similar to

the *E. coli* -1 product (Figure 4D). A -1 frame-shift would encode a protein with a 12-amino acid extension. Such an extension could permit an interaction that may further distinguish γ from τ functionally or could loop back to stabilize its structure in a thermal environment. Resolution of this issue will require further investigation since the two alternatives differ by ca 1 kDa, less than the resolution of SDS PAGE.

To date only one DNA polymerase has been discovered in thermophilic eubacteria, even though others have been actively sought (Chien *et al.*, 1976; Kaledin *et al.*, 1980; Lawyer *et al.*, 1989). Our studies establish that one extreme thermophile has a Pol III-class polymerase and an associated Pol III holoenzyme auxiliary subunit homolog. This provides strong evidence for the existence of the standard Pol III holoenzyme-replication mechanism in this one thermophile and establishes a principle that could be extended to other thermophiles. Since *T. thermophilus* (and *T. aquaticus*) DNA polymerase is homologous to DNA polymerase I (Lawyer *et al.*, 1989) and since we have established the existence of an additional polymerase in this organism, we propose the enzyme be referred to as *T. thermophilus* DNA polymerase I in the future to distinguish the two enzymes.

E. coli Pol III holoenzyme can remain associated with primed DNA for 40 minutes and replicates DNA at 500 to 1000 nucleotides/second (McHenry, 1988). Existing PCR technology is limited by relatively non-processive repair-like DNA polymerases. Our studies reveal the existence of a thermophilic replicase that, by inference, likely processes rapid replication and highly processive properties at elevated temperatures. It is likely that such an enzyme can be developed into a tool for megabase PCR, removing the current length restrictions and enabling new technological advances in molecular biology.

Methods

General methods

Protein determinations were conducted using the Coomassie Protein Assay Reagent from Pierce and BSA as a standard. DNA manipulations, cloning, Southern Blots, and standard molecular biological methods were carried out as described (Ausubel *et al.*, 1995). Conductivity measurements were performed using a Radiometer CDM83 conductivity meter on samples diluted 1:100 in distilled water equilibrated to room temperature. Standards were run in parallel.

Identification of monoclonal antibodies that cross-react with *T. thermophilus* α subunit of Pol III

In general, large asymmetric complexes are relatively insoluble in ammonium sulfate. Ammonium sulfate fractionation provides a 50-

fold purification of the *E. coli* Pol III holoenzyme. We took a low ammonium sulfate cut of *T. thermophilus* extracts to provide a source of protein enriched sufficiently in Pol III that we could use an ELISA assay to screen 12 monoclonal antibodies directed against the *E. coli* Pol III α subunit (M. Olson & C. S. M., unpublished) to see if they cross-reacted with a *T. thermophilus* protein. The ammonium sulfate fraction was prepared by addition of 0.246 g ammonium sulfate to each ml of Fr I using the same approach as described for the preparation of Fr II in the Pol III preparation described below. For the ELISA screening assay, all manipulations were conducted at room temperature. Into each well of a 96-well microtiter plate (Corning Costar High Binding EIA/RIA) was placed 4 μ g protein in 150 μ l Buffer E7 (10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween-20). After an overnight incubation, each well was blocked by incubation with buffer E7 + 10 μ g/ml BSA for three hours, followed by three washes with buffer E7 + 10 mg/ml BSA and the addition of 150 μ l of hybridoma supernatant and a 3.5 hours incubation. Wells were then washed three times with buffer E7 and once with an equivalent buffer that had been adjusted to pH 8.8 (buffer E8). An amount (150 μ l) of a 1:3000 dilution of goat anti-mouse IgG antibody-alkaline phosphatase conjugate (BioRad) was added and incubated for one hour. Wells were washed three times with buffer E7, once with buffer E8 and developed in the presence of *p*-nitrophenyl phosphate for ten minutes as described (Harlow & Lane, 1988). Absorbance was read at 405 nm. Monoclonal antibodies produced from hybridoma lines C1950-F3 and C1104-H2 gave an absorbance of 0.17 and 0.16, respectively; all other candidate monoclonal supernatants gave an absorbance of 0.06 on average. We combined equal volumes of both supernatants for future work. An amount (40 μ g) of the 0.246 ammonium sulfate fraction of *T. thermophilus* Fr I was subjected to the Western blotting procedure described below. A band migrating with the same mobility as the α subunit of Pol III of *E. coli* was detected. Control experiments using several different monoclonal antibodies did not result in detection of the same band.

Gap filling polymerase assay

The gap filling assay used, is the same as that used to monitor the purification of *E. coli* Pol III core (Kim & McHenry, 1996) except that the assay mixture contained 32 mM Hepes (pH 7.5), 13% glycerol, 0.01% Nonidet P40, 0.13 mg/ml BSA, 10 mM $MgCl_2$, 0.2 mg/ml activated calf-thymus DNA, 57 μ M each of dGTP, dATP, dCTP and 21 μ M [3H]dTTP (360 cpm/pmol dTTP). The reaction was started by the addition of polymerase in a 15 μ l reaction mixture and incubated at 37°C for five to

ten minutes. Acid insoluble DNA product was recovered and quantified as described (Kim & McHenry, 1996). The assay was performed at 37°C, in part for convenience, but also because the decrease in activity for Pol III at lower temperatures is less than that of the major *T. thermophilus* DNA polymerase activity, permitting more efficient detection of Pol III because it comprised a greater percentage of total polymerase activity when assayed this way. One unit of activity is equivalent to 1 pmol total nucleotide synthesis per minute at 37°C. The ratio of activity for Fr V *T. thermophilus* Pol III between 37°C and 60°C is five-fold.

Western blotting procedures

Aliquots of column fractions (50 µl) were subjected to electrophoresis on a 10% SDS/polyacrylamide gel. Proteins were then transferred from the gel to a PVDF membrane (BioRad). Transfer buffer contained 25 mM Tris, 192 mM glycine, 20% methanol and was adjusted to pH 8.5 by the addition of HCl. Transfer was conducted for two hours at 70 V (0.7 A). The membrane was then washed in 10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween-20 followed by incubation for one hour in the same buffer with 5% dried milk. The blot was then incubated overnight with hybridoma supernatant, washed (three times) in 100 ml buffer TBS containing 0.5% Tween-20, incubated for one hour with a 1:2000 dilution of goat anti-mouse IgG alkaline phosphatase conjugate (BioRad) in buffer TBS containing 0.5% dried milk, washed (three times) in 100 ml buffer TBS + 0.5% Tween-20 and once in 100 ml buffer P (100 mM Tris-HCl (pH 9.5), 100 mM NaCl, 5 mM MgCl₂). The membrane was developed by incubating in a 1:150 dilution of a GIBCO nitroblue tetrazolium stock solution and a 1:300 dilution of a 5-bromo-4-chloro-3-indoyl phosphate stock solution in buffer B until the intensity of the α band reached the desired intensity. The development of color was stopped by washing the membrane in water.

Partial purification of *T. thermophilus* Pol III α subunit and associated proteins.

T. thermophilus strain pMF48KAT (Lasa *et al.*, 1992) was grown in 180 liter batches with aeration in a 250 liter fermentor at 72°C in a medium containing (per liter): 0.27 g ferric chloride hexahydrate, 0.294 g sodium citrate trisodium salt dihydrate, 0.025 g calcium sulfate dihydrate, 0.20 g magnesium chloride hexahydrate, 0.53 g ammonium chloride, 8 g pancreatic digest of casein (DMV International), 4 g yeast extract (Ardamine Z, Champlain Industries), 4.0 g glucose (Cerelose 2001, food grade (Corn Products, International), 0.5 g L-glutamic acid monosodium salt, 0.254 g sodium phosphate monobasic (dissolved in 5.6 ml water), 1.5 g dipotassium phosphate (dissolved in

11.1 ml water). The medium was maintained at pH 7.3 by the addition of 3 M NaOH. An additional five liter solution of glucose (480 g/l) and L-glutamic acid (60 g/l) was added incrementally during the course of the fermentation run. A 250 ml overnight inoculum (grown in the same medium + 30 µg/ml kanamycin) was used. Cells were harvested by passage through a heat exchanger in a Sharples AS 16 centrifuge and resuspended in an equal weight of Tris-sucrose (50 mM Tris-HCl (pH 7.5), 10% sucrose) frozen in liquid N₂ and stored at -20°C until used.

4.7 kg of a 1:1 suspension of cells in Tris-sucrose were added to 6.5 liters Tris-sucrose that had been prewarmed to 55°C. To the stirred mixture was added 117 ml 0.5 M DTT and 590 ml of 2 M NaCl, 0.3 M spermidine in Tris-sucrose adjusted to pH 7.5. The pH of the slurry was adjusted to pH 8 by the addition of 2 M Tris base and 2.35 g lysozyme was added. The slurry was distributed into 250 ml centrifuge bottles and incubated at 30°C for one hour with occasional inversion and then centrifuged at 23,000 g for 60 minutes at 4°C. The recovered supernatant (eight liters) constituted Fr I (Table 1).

To Fr I was added ammonium sulfate (0.267 g/initial ml Fr I) over a 15 minute interval. The mixture was stirred for an additional 30 minutes at 4°C and then centrifuged at 23,000 g for 60 minutes at 0°C. The recovered pellet was resuspended (on ice) in 563 ml (0.07 \times Fr I volume), 50 mM Tris-HCl (pH 7.5), 20% glycerol, 1 mM EDTA, 0.1 M NaCl, 5 mM DTT, 0.18 g (added to each ml of final solution) ammonium sulfate and the resulting suspension centrifuged at 23,000 g for 60 minutes at 4°C resulting in Fr II (Table 1).

A BioRex 70 cation exchange chromatography step was developed that resolved polymerase activity from the majority of contaminating protein. Fr II was resuspended in 100 ml buffer U (50 mM Imidazole-HCl (pH 6.8), 20% glycerol, 35 mM ammonium sulfate, 1 mM magnesium acetate, 0.1 mM zinc sulfate, 5 mM β -mercaptoethanol, 0.1 mM ATP) and dialyzed twice successively versus two liters buffer U. Dialysate was applied to a 300 ml BioRex 70 (BioRad, 100 to 200 mesh, 5.5 cm diameter) column equilibrated in buffer U and washed with 0.9 liter buffer U. Activity was eluted with a 1.5 liter 0 to 300 mM NaCl gradient in buffer U. All gradient eluted fractions containing greater than 20,000 units of gap-filling polymerase activity/ml were pooled, constituting Fr. III (Table 1).

Hydrophobic interaction chromatography was used to resolve a unique polymerase that cross-reacted with antibody directed against *E. coli* Pol III α subunit. Fr III protein was precipitated by the addition of an equal volume of saturated ammonium sulfate and the pellet was collected by centrifugation at 23,000 g, one hour, 4°C. The pellet was dissolved in 40 ml buffer U and 20 ml Taya-Pearl-Ether 650 M (Toso Haas) equilibrated in buffer U + 1.6 M ammonium sulfate was added. To

the stirred suspension was added (dropwise) 33.6 ml 4 M ammonium sulfate and the entire mixture was applied to an 80 ml ToyoPearl-Ether column (3.5 cm diameter) equilibrated in buffer U + 1.6 M ammonium sulfate. The column was washed with 80 ml buffer U + 1.1 M ammonium sulfate and polymerase activity eluted with a 1.5 liter 1.1 to 0.3 M ammonium sulfate gradient in buffer U (Figure 1). Fractions containing polymerase activity were subjected to the Western blotting procedure. The second peak (fractions 45 to 50) that contained polymerase activity and reacted with monoclonal antibodies against *E. coli* Pol III α subunit were pooled, constituting Fr IV (120 ml, Table 1, Figure 1A,B).

One-half of Fr IV was combined with an equal volume of saturated ammonium sulfate (4°C) and the resulting precipitate collected by centrifugation and redissolved in 4 ml buffer U and dialyzed overnight versus 250 ml buffer U (one buffer change) and the dialysate was applied to a 5 ml Q-Sepharose fast flow column (1.4 cm diameter) equilibrated in buffer U. The column was washed with 15 ml buffer U and polymerase activity was eluted with a 75 ml 0 to 275 mM NaCl gradient in buffer U. A contaminating protein eluted toward the end of the activity peak; thus, fractions were carefully selected for pooling based on purity. All fractions that contained greater than 50 μ g protein and a specific activity greater than 104,000 were pooled (Fr. 28 to 36) to yield Fr V (25 ml, Table 1, Figure 2).

Procedures used in obtaining protein sequence

N-terminal sequencing was performed on the τ (63 kDa) and γ (50 kDa) proteins (Fr. V, 120 μ g total protein) separated on a 10% SDS polyacrylamide gel and transferred to a Hyperbond (BioRad) PVDF membrane (Matsudaira, 1987). The appropriate Coomassie blue stained band was cut out and sequenced on an ABS 477A Protein Sequencer according to manufacturer's instructions. N-terminal sequencing was performed by Dr James McManaman in the UCHSC Cancer Center Protein microsequencing facility.

Internal peptide sequences were determined at the Harvard Microchemistry facility directed by Dr Wm Lane. Their standard procedures were followed. An aliquot of Fr. V (230 μ g) was subjected to electrophoresis on a 10% polyacrylamide gel and the fractionated proteins transferred to a PVDF membrane (BioRad). The Ponceau-S stained 63 kDa band was cut out and subjected to digestion by endo Lys-C, the resultant peptide separated by HPLC and one peptide was chosen for sequencing.

Isolation of *T. thermophilus* dnaX probe by PCR

Forward primers were selected from the amino-terminal protein sequence that would provide the least degeneracy; nevertheless, two separate primers had to be synthesized so that no primer was

more than 512-fold degenerate. Primer X1Fa (5'-TTY CAR GAR GTN GTN GGW CA) and Primer X1Fb (5'-TTY CAR GAR GTN GTN GGS CA) differed only in the sixth codon. For the reverse primer, Primer X139R was used (5'-GGY TCY TCN ARN GTY TT; 256-fold degenerate). For PCR reactions, the Boehringer Mannheim Expand™ long template PCR system was used and we followed the manufacturer's recommendations except as noted. Reactions were conducted in Boehringer Expand™ buffer 1 supplemented with 0.5 mM extra Mg^{2+} . The annealing steps were conducted at 48°C, elongation at 68°C and the melting step at 94°C; 26 total cycles were run. Products were separated on 2% FMC Metaphor agarose gels, visualized by Sybr Green-I staining, extracted and cloned into vector pCRII (Invitrogen). A plasmid isolate containing an *Eco*RI fragment of the predicted size based upon the distance between the homologous region of the *E. coli* *dnaX* gene (393 bases) was submitted for DNA sequencing and was shown to encode a 388 bp segment highly homologous to other bacterial *dnaX* genes (42% identical to *B. subtilis* *dnaX* in the segment between the regions corresponding to the primers).

Cloning and sequencing of full-length *T. thermophilus* dnaX gene

A *Pst*I digest of *T. thermophilus* chromosomal DNA was subjected to electrophoretic separation on an agarose gel and the region corresponding to a 7.1 kb fragment that hybridized with the partial *dnaX* probe generated by PCR was extracted and cloned into vector pMGC707. pMGC707 had been prepared by cutting pCRII (Invitrogen) with *Spe*I and *Not*I and inserting a polylinker resulting from the annealing oligonucleotides (5'-GGCCGC-AATTGCACGCGTTCGAATTCATGACGTCTTC-CAGTGCACCTGGTTAATTA) and (5'-CTAGTTA-ATTAACCAGTGCACCTGGAAGACGTCATGG-AATTCGAACGCGTGCATTTG). The polylinker region of the resulting plasmid was cleaved with *Bst*XI to generate *Pst*I compatible termini to receive the extracted 7.1 kb population of fragments. Colonies were screened by colony hybridization (Ausubel *et al.*, 1995) using the *T. thermophilus* *dnaX* PCR-generated probe. Plasmids from positive colonies were purified and those containing 7.1 kb inserts and also a *Bam*HI site as indicated from the sequence of the PCR probe were retained for further characterization. One was submitted for full DNA sequencing (Lark Technologies, Houston, TX), resulting in the sequence of the full length gene (Figure 4).

Acknowledgments

We are indebted to Michael Fulton for growing the *T. thermophilus* cells used in this study in the UCHSC Fermentation Facility and for his assistance on some of

the purification steps. This work was supported by a contract to the UCHSC from Enzyco and by an NIH SBIR grant to Enzyco.

References

- Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J. & Struhl, K. (1995). *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., New York.
- Blinkowa, A. & Walker, J. (1990). Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the τ subunit reading frame. *Nucl. Acids Res.* **18**, 1725–1729.
- Brush, G., Kelly, T. & Stillman, B. (1995). Identification of eukaryotic DNA replication proteins using Simian Virus 40 *in vitro* replication system. *Methods Enzymol.* **262**, 522–548.
- Carter, J., Franden, M., Aebersold, R. & McHenry, C. (1993). Identification, isolation and characterization of the structural gene encoding the δ' subunit of *E. coli* DNA polymerase III holoenzyme. *J. Bacteriol.* **175**, 3812–3822.
- Chien, A., Edgar, D. & Trela, J. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* **127**, 1550–1557.
- Cull, M. & McHenry, C. (1995). Purification of *E. coli* DNA polymerase III holoenzyme. *Methods Enzymol.* **262**, 22–35.
- Dallmann, H. G. & McHenry, C. (1995). DnaX-complex of *Escherichia coli* DNA polymerase III holoenzyme: physical characterization of the DnaX subunits and complexes. *J. Biol. Chem.* **270**, 29563–29569.
- Dallmann, H. G., Thimmig, R. & McHenry, C. (1995). DnaX-complex of *E. coli* DNA polymerase III holoenzyme: central role of τ in initiation complex assembly and in determining the functional asymmetry of holoenzyme. *J. Biol. Chem.* **270**, 29555–29562.
- Flower, A. & McHenry, C. (1990). The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proc. Natl Acad. Sci. USA*, **87**, 3713–3717.
- Hammond, R. & Brown, N. (1992). Overproduction and purification of *Bacillus subtilis* DNA polymerase III. *Protein Expression Purif.* **3**, 65–70.
- Harlow, E. & Lane, D. (1988). In *Antibodies. A Laboratory Manual*, p. 257, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Kaledin, A., Slyusarenko, A. & Gorodetskii, S. (1980). Vydelenie i svoistva DNK-polimerazy iz ekstremal'noi termofil'noi bakterii *Thermus aquaticus* YT1. *Bio-khimiya*, **45**, 644–651.
- Kim, D. R. & McHenry, C. (1996). *In vivo* assembly of overproduced DNA polymerase III: overproduction, purification and characterization of the α , α - ϵ and α - ϵ - θ subunits. *J. Biol. Chem.* **271**, 20681–20689.
- Kong, X., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992). Three-dimensional structure of the β subunit of *Escherichia coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.
- Kornberg, A. & Baker, T. (1992). *DNA Replication*, 2nd edit., W. H. Freeman and Company, New York.
- Kornberg, T. & Geffer, M. (1972). Deoxyribonucleic acid synthesis in cell-free extracts. *J. Biol. Chem.* **247**, 5369–5375.
- LaDuca, R., Crute, J., McHenry, C. & Bambara, R. (1986). The β subunit of the *Escherichia coli* DNA polymerase III holoenzyme interacts functionally with the catalytic core in the absence of other subunits. *J. Biol. Chem.* **261**, 7550–7557.
- Lasa, I., Castón, J., Fernández-Herrero, L., de Pedro, M. & Berenguer, J. (1992). Insertional mutagenesis in the extreme thermophilic eubacteria *Thermus thermophilus* HB8. *Mol. Microbiol.* **6**, 1555–1564.
- Lawyer, F., Stoffel, S., Saiki, R., Myambo, K., Drummond, R. & Gelfand, D. (1989). Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J. Biol. Chem.* **264**, 6427–6437.
- Low, R., Rashbaum, S. & Cozzarelli, N. (1976). Purification and characterization of DNA polymerase III from *Bacillus subtilis*. *J. Biol. Chem.* **251**, 1311–1325.
- Matsudaira, P. (1987). Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* **262**, 10035–10038.
- McHenry, C. (1982). Purification and characterization of *Escherichia coli* DNA polymerase III. Identification of τ as a subunit of the DNA polymerase III holoenzyme. *J. Biol. Chem.* **257**, 2657–2663.
- McHenry, C. (1988). DNA polymerase III holoenzyme of *Escherichia coli*. *Annu. Rev. Biochem.* **57**, 519–550.
- McHenry, C. & Crow, W. (1979). DNA polymerase III of *Escherichia coli*: purification and identification of subunits. *J. Biol. Chem.* **254**, 1748–1753.
- McHenry, C. & Kornberg, A. (1977). DNA polymerase III holoenzyme of *Escherichia coli*: purification and resolution into subunits. *J. Biol. Chem.* **252**, 6478–6484 and 253, 645.
- Mok, M. & Marians, K. (1987). The *Escherichia coli* pre-primosome and dnaB helicase can form replication forks that move at the same rate. *J. Biol. Chem.* **262**, 16644–16654.
- Naktinis, V., Onrust, R., Fang, L. & O'Donnell, M. (1995). Assembly of a chromosomal replication machine: Two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. II. Intermediate complex between the clamp loader and its clamp. *J. Biol. Chem.* **270**, 13358–13365.
- O'Donnell, M., Onrust, R., Dean, F., Chen, M. & Hurwitz, J. (1993). Homology in accessory proteins of replicative polymerases: *E. coli* to humans. *Nucl. Acids Res.* **21**, 1–3.
- Sherman, F., Stewart, J. & Tsunasawa, S. (1985). Methionine or not methionine at the beginning of a protein. *Bioessays*, **3**, 27–31.
- Studwell-Vaughan, P. & O'Donnell, M. (1991). Constitution of the twin polymerase of DNA polymerase III holoenzyme. *J. Biol. Chem.* **266**, 19833–19841.
- Tsuchihashi, Z. & Kornberg, A. (1990). Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. *Proc. Natl Acad. Sci. USA*, **87**, 2516–2520.
- Uhlmann, F., Cai, J., Flores-Rozas, H., Dean, F., Finkelstein, J., O'Donnell, M. & Hurwitz, J. (1996). *In vitro* reconstitution of human replication factor C from its five subunits. *Proc. Natl Acad. Sci. USA*, **93**, 6521–6526.
- Wickner, S. (1976). Mechanism of DNA elongation catalyzed by *Escherichia coli* DNA polymerase III, dnaZ protein and DNA elongation factor I and factor III. *Proc. Natl Acad. Sci. USA*, **73**, 3511–3515.

- Wu, C., Zechner, E. & Mariani, K. (1992). Coordinated leading and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J. Biol. Chem.* **267**, 4030–4044.
- Wu, Y., Franden, M., Hawker, J. & McHenry, C. (1984). Monoclonal antibodies specific for the α subunit of the *Escherichia coli* DNA polymerase III holoenzyme. *J. Biol. Chem.* **259**, 12117–12122.
- Yoder, B. & Burgers, P. (1991). *Saccharomyces cerevisiae* replication factor C I purification and characterization of its ATPase activity. *J. Biol. Chem.* **266**, 22689–22697.

Edited by A. R. Fersht

(Received 15 May 1997; received in revised form 27 June 1997; accepted 27 June 1997)

[EBI Home](#)[About EBI](#)[Research](#)[Services](#)[Toolbox](#)[Databases](#)[Downloads](#)

SEQUENCE ANALYSIS

- [Help](#)
- [General Help](#)
- [Formats](#)
- [Gaps](#)
- [Matrix](#)
- [References](#)
- [ClustalW Help](#)
- [ClustalW FAQ](#)
- [Jalview Help](#)
- [Scores Table](#)
- [Alignment](#)
- [Guide Tree](#)
- [Colours](#)

ClustalW Results

Results of search	
Number of sequences	2
Alignment score	16190
Sequence format	Pearson
Sequence type	nt
ClustalW version	1.82
JalView	
Output file	clustalw-20050410-16395846.output
Alignment file	clustalw-20050410-16395846.aln
Guide tree file	clustalw-20050410-16395846.dnd
Your input file	clustalw-20050410-16395846.input
SUBMIT ANOTHER JOB	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

[View Output File](#)

SeqA Name	Len (nt)	SeqB Name	Len (nt)	Score
1 McHenry	2681	2 Yurieva	2007	98

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple sequence alignment.

[View Output File](#)

Alignment

[Show Colors](#)[View Alignment File](#)

CLUSTAL W (1.82) multiple sequence alignment

```
McHenry      TCCGCCAGGCGCTTTCCCGTGAGGTAGAGGGCCATCTCCACCACCTCGAGCT
Yurieva      -----

McHenry      GGCGTCACCCGAAGCCCCCCCACGAAGCGGCTCTCGTACCCAGGCGCCCA
Yurieva      -----

McHenry      CCGATCTCGGGCCCCGCCCCCGTGGACGAGAACCAAGGGACCAGGCTAGGCGG
Yurieva      -----

McHenry      TCCAATAGGGCCTCCGCCCCCTAAGGCTTCCTCCCACCTTCACCAAAAGGG
Yurieva      -----

McHenry      AAGGTAGACCACCTCCTCGGGCAGGCCCTCCTCCTCCTCGGCCTCCAGGAG
Yurieva      -----

McHenry      GCCCAGAAGGGCGTGGTCCGGGTGGAGGCCGAAGCGGGCGGCGAGGGCCCGG
Yurieva      -----

McHenry      CTCGGGCGGGAGGGCCTGGGCCCTCGTCAGGAGGGGGCGAGGTCGGCCGGG
Yurieva      -----

McHenry      GGCCGTCCCAAGCTCCGGACCCGCGCGAAGCTCCGCCCTCACCTCCCCCCCC
Yurieva      -----

McHenry      GAGGAAGAGGAGGTCCTCCTCGCTTAGTACCAGGAAGGCCAAGGCCTGCCCT
Yurieva      -----

McHenry      AAGCTCCCGCAGGAAGGCCCGGATGGCCTCCGGGTCTCCTCGCCTCGAGG
Yurieva      -----

McHenry      GTAGAGGCTCGTCCAAGGGGGGTGCGGGACCAGGTAGACCCCGTCCCTCCC
Yurieva      -----TCCGGGGGTGG--GGTCCCAGGTAGACCCGCCCCCTCCC
          ***  ***  **  *  *****  *  *****

McHenry      CTCGGCCCAGGCCGCCACCGCCTCCAGGGGGGCTGCACGTGCAAGGAGAGG
Yurieva      CTTTACCCAGGCCGCCACCTCCTCCAGGGGGGCAAGGCGTCAAGGAGAGG
          **  *****  *****  *****

McHenry      CACCACGCCCTATACTAGCCCTTGAGAGCGCCCTCTACCGCGCTTCCGCCC
Yurieva      CACCACGCCCTATACTAGCC--TTGTGAGCGCCCTCTACCGCGCTTCCGCCC
          *****  *****  *****  *****

McHenry      TCCAGGAGGTGGTGGGGCAGGAGCACGTGAAGGAGCCCCCTCCTCAAGGCCAT
Yurieva      TCCAGGAGGTGGTGGGGCAGGAGCACGTGAAGGAGCCCCCTCCTCAAGGCCAT
          *****  *****  *****  *****

McHenry      GGAGGCTCGCCAGGCCTACCTCTTCTCCGGGCCCAGGGGCGTGGGCAAGAC
```

```
Yurieva      GGAGGCTCGCCAGGCCTACCTCTTCTCCGGGCCAGGGGCGTGGGCAAGAC(
*****

McHenry
Yurieva      CGAGGCTCCTCGCCATGGCGGTGGGGTGCCAGGGGGAAGACCCCCCTTGCGG(
CGAGGCTCCTCGCCATGGCGGTGGGGTGCCAGGGGGAAGACCCCCCTTGCGG(
*****

McHenry
Yurieva      CCCACTGCCAGGCGGTGCAGAGGGGCGCCACCCGGACGTGGTGGAGATTGA(
CCCACTGCCAGGCGGTGCAGAGGGGCGCCACCCGGACGTGGTGGACATTGA(
*****

McHenry
Yurieva      GCAACAACCTCCGTGGAGGACGTGCGGGAGCTGAGGGAAAGGATCCACCTCGC(
GCAACAACCTCCGTGGAGGACGTGCGGGAGCTGAGGGAAAGGATCCACCTCGC(
*****

McHenry
Yurieva      CTGCCCCAGGAAGGTCTTCATCCTGGACGAGGCCACATGCTCTCCAAAAG(
CTGCCCCAGGAAGGTCTTCATCCTGGACGAGGCCACATGCTCTCCAAAAG(
*****

McHenry
Yurieva      ACGCCCTCCTCAAGACCCTGGAGGAGCCCCGCCCCACGTCTTCTCGTCTT(
ACGCCCTCCTCAAGACCCTGGAGGAGCCCCGCCCCACGTCTTCTCGTCTT(
*****

McHenry
Yurieva      CCGAGCCCGAGAGGATGCCCCCACCATCCTCTCCCGCACCAGCACTTCCG(
CCGAGCCCGAGAGGATGCCCCCACCATCCTCTCCCGCACCAGCACTTCCG(
*****

McHenry
Yurieva      GCCTCACGGAGGAGGAGATCGCCTTTAAGCTCCGGCGCATCCTGGAGGCCGT(
GCCTCACGGAGGAGGAGATCGCCTTTAAGCTCCGGCGCATCCTGGAGGCCGT(
*****

McHenry
Yurieva      AGGCGGAGGAGGAGGCCCTCCTCCTCCTCGCCCGCTGGCGGACGGGGCCCT(
AGGCGGAGGAGGAGGCCCTCCTCCTCCTCGCCCGCTGGCGGACGGGGCCCT(
*****

McHenry
Yurieva      CGGAAAGCCTCCTGGAGCGCTTCTCCTCCTGGAAGGCCCCCTCACCCGGA(
CGGAAAGCCTCCTGGAGCGCTTCTCCTCCTGGAAGGCCCCCTCACCCGGA(
*****

McHenry
Yurieva      AGCGCGCCCTAGGCCTCCCCCAGGGAGGCCCTGGCCGAGATCGCCGCCTC(
AGCGCGCCCTAGGCCTCCCCCAGGGACCGGGTGGCCGAGATCGCCGCCTC(
*****

McHenry
Yurieva      GGGGGAAAACGGCGGAGGCCCTGGGCCTCGCCCGCGCCTCTACGGGGAAGG(
GGGGGAAAACGGCGGAGGCCCTGGGCCTCGCCCGCGCCTCTACGGGGAAGG(
*****

McHenry
Yurieva      CGAGGAGCCTGGTCTCGGGCCTTTTGAGGTGTTCCGGGAAGGCCTCTACGC(
CGAGGAGCCTGGTCTCGGGCCTTTTGAGGTGTTCCGGGAAGGCCTCTACGC(
*****

McHenry
Yurieva      GCCTCGCGGGAACCCCCCTTCCCGCCCCGCCCCAGGCCCTGATCGCCGCCAT(
GCCTCGCGGGAACCCCCCTTCCCGCCCCGCCCCAGGCCCTGATCGCCGCCAT(
*****

McHenry
Yurieva      TGGACGAGGCCATGGAGCGCTCGCCCGCCGCTCCGACGCCTTAAGCCTGGA(
TGGACGAGGCCATGGAGCGCTCGCCCGCCGCTCCGACGCCTTAAGCCTGGA(
*****

McHenry
Yurieva      TCCTGGAGGCGGGAAGGGCCCTGGCCGCGGAGGCCCTGCCAGCCACGGG(
TCCTGGAGGCGGGAAGGGCCCTGGCCGCGGAGGCCCTACCCAGCCACGGG(
*****

McHenry
Yurieva      CCCCAGAGGTGCGCCCCAAGCCGGAAGCCCCCGGCCCGGAACCCCCAAG(
CCCCAGAGGTGCGCCCCAAGCCGGAAGCCCCCGGCCCGGAACCCCCAAG(
*****
```

```
McHenry      AGGCGCCCGACCTGCGGGAGCGGTGGCGGGCCTTCCTCGAGGCCCTCAGGCC(
Yurieva      AGGCGCCCGACCTGCGGGAGCGGTGGCGGGCCTTCCTCGAGGCCCTCAGGCC(
*****

McHenry      GGGCCTTCGTGCGGGAGGCCCGCCCGGAGGTCCGGGAAGGCCAGCTCTGCCT(
Yurieva      GGGCCTTCGTGCGGGAGGCCCGCCCGGAGGTCCGGGAAGGCCAGCTCTGCCT(
*****

McHenry      CCGAGGACAAGGCCTTCCACTACCGCAAGGCCTCGGAACAGAAGGCGAGGCT(
Yurieva      CCGAGGACAAGGCCTTCCACTACCGCAAGGCCTCGGAACAGAAGGTGAGGCT(
*****

McHenry      TGGCCCAGGCCCATTTCGGGGTGGAGGAGGTGTTCTCGTCTTGAGGGGAGA(
Yurieva      TGGCCCAGGCCCATTTCGGGGTGGAGGAGGTGTTCTCGTCTTGAGGGGAGA(
*****

McHenry      GCCTGAGCCCAAGGCCCCGCTCGGCCCCACCTCCTGAAGCGCCCGCACCCT(
Yurieva      GCCTGAGCCCAAGGCCCCGCTCGGCCCCACCTCCTGAAGCGCCCGCACCCT(
*****

McHenry      CCGAGGAGGAGGTAGAGGCGGAGGAAGCGGCGGAGGAGGCCCGGAGGAGGC(
Yurieva      CCGAGGAGGAGGTAGAGGCGGAGGAAGCGGCGGAGGAGGCCCGGAGGAGGC(
*****

McHenry      GGGTGGTCCGCCTCCTGGGGGGGCGGGTGCTCTGGGTGCGGCGGCCAGGAC(
Yurieva      GGGTGGTCCGCCTCCTGGGGGGGCGGGTGCTCTGGGTGCGGCGGCCAGGAC(
*****

McHenry      CGCCGGAGGAGGAACCCCTGAGCCAAGACGAGATAGGGGGTACTGGTATATA(
Yurieva      CGCCGGAGGAGGAACCCCTGAGCCAAGACGAGATAGGGGGTACTGGTATATA(
*****

McHenry      ATGACGCGGACCACCGACCTCGGACAAGAGACCGTGGAACAACATCCTCAAGC(
Yurieva      ATGACGCGGACCACCGACCTCGGACAAGAGACCGTGGAACAACATCCTCAAGC(
*****

McHenry      CGTATTGAGGGCCAGGTGCGGGGGCTCCAAAAGATGGTGCCGAGGGCCGCC(
Yurieva      CGTATTGAGGGCCAGGTGCGGGGGCTCCAGAAGATGGTGCCGAGGGCCGCC(
*****

McHenry      GAGGTCCTCAGCCAGATGACCGCCACCAAGAAGGCCATGGAGGCGGCGGCCA(
Yurieva      GAGGTCCTCAGCCAGATGACCGCCACCAAGAAGGCCATGGAGGCGGCGGCCA(
*****

McHenry      CTCCACGAGTTCCTGAACGTCTGCGCCGCGGAGGTCTCCGAGGGCAAGGTGA(
Yurieva      CTCCACGAGTTCCTGAACGTCTGCGCCGCGGAGGTCTCCGAGGGCAAGGTGA(
*****

McHenry      AAGCCCGAGGAGATCGCCACCATGCTGAAGAAGTTCATCTAGATGGGTGGCC(
Yurieva      AAGCCCGAGGAGATCGCCACCATGCTGAAGAAGTTCATCTAGATGGGTGGCC(
*****

McHenry      CGCCTCCGGCGCCTCCTCCGGGCCCTTCTCGCCAGGAGGC 2681
Yurieva      -----
```

PLEASE NOTE: Showing colors on large alignments is slow.

Show Colors

View Alignment File

Guide Tree

[Show as Phylogram Tree](#)[Show Distances](#)[View DND File](#)

(McHenry:0.00897,Yurieva:0.00897);

Cladogram

[Show as Phylogram Tree](#)[Show Distances](#)[View DND File](#)

Right-click on the above tree to see display options.

Problems printing? Read [how to print a Phylogram or Cladogram](#).

Please contact [EBI Support](#) with any problems or suggestions regarding this site.

 [View Printer-friendly version of this page](#) | [Terms of Use](#)

ClustalW Results

Results of search	
Number of sequences	2
Alignment score	2701
Sequence format	Pearson
Sequence type	aa
ClustalW version	1.82
JalView	
Output file	clustalw-20050511-23053240.output
Alignment file	clustalw-20050511-23053240.aln
Guide tree file	clustalw-20050511-23053240.dnd
Your input file	clustalw-20050511-23053240.input
<input type="button" value="SUBMIT ANOTHER JOB"/>	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

SeqA Name	Len(aa)	SeqB Name	Len(aa)	Score
=====	=====	=====	=====	=====
1 McHenry	528	2 Yurieva	497	98
=====	=====	=====	=====	=====

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Please check the out

Alignment

CLUSTAL W (1.82) multiple sequence alignment

```

McHenry      -SALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 59
Yurieva      MSALYRRFRPLTFQEVG-QEHVKEPLLKAIRE-RLAQAYLFSGPRGVG-TTTARLLAMAV 57
*****

```

McHenry	GCQGEDPPCGVCPHCQAVQGAHPDVVEIDAASNNSVEDVRELRLRIHLAPLSAPRKVFI	119
Yurieva	GCQG-DPPCGVCPHCQAVQR-AHPDVVDIDAASNNS-EDVRELRLRIHLAPL-APRKVFI	113

McHenry	LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTQHFRRRLTEEEIA	179
Yurieva	LDEAHMLS-SAFNALLKTLLEPPP-VLFVFATTEPERMPP-ILSRTQHFRRRLT-EEIA	169

McHenry	FKLRRILEAVGREAEAEALLLRLADGALRDAESLLERFLLLEGPLTRKEVERALGLPP	239
Yurieva	FKLRRILEAVG-RAEEAEALLLRLADG-LRDAESLLERFLLLE-PLTRKEVERALGSPP	226

McHenry	REALAEIAASLARGKTAEALGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTPLP	299
Yurieva	-TGVAEIAASLARGKT-EALGLARRLYGEGYA-RSLVSGLLEVFREGL-AAFGLAGTPLP	282

McHenry	APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPPPEVGPKP	359
Yurieva	APPQ-LIAAMTALDEAMERLRR-SDALSLEVALLEA-RALAAEALPQPTGAPP-EVGPKP	338

McHenry	ESPPAPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY	419
Yurieva	ESPPTPEP-RPEEAPDLRERWRAF-EALRPTLRAFVREAR-EVREGQLCLAFPEDK-FHY	394

McHenry	RKASEQKARLLPLAQAHFGVEEVVLVLEGEKKSLSPRPRPAPPPEAPAPPGPPEEEVEAE	479
Yurieva	RKASEQKVRLLP-AQAHFGVEEVVLVLE-EKKSLSPRPRPAPP-APAPPGPPEEEVEA-	450

McHenry	EAAEEAPEEALRRVRLGGRVLWVRRPRTREAPPEEPLSQDEIGGTGI	528
Yurieva	EAAEEAPEEALRRVVR-LGGRVLWVRRPRTRE-PEEPLSQDEIGGTGI	497

PLEASE NOTE: Showing colors on large alignments is slow.

Show Colors

View Alignment File

Guide Tree

Show as Phylogram Tree

Show Distances

View DND File

(McHenry:0.00905,Yurieva:0.00905);

Cladogram



McHe
Yuriev

Show as Phylogram Tree

Show Distances

View DND File

Right-click on the above tree to see display options.

Problems printing? Read [how to print a Phylogram or Cladogram](#).

[Return to site](#)

ClustalW Results

Results of search	
Number of sequences	2
Alignment score	2276
Sequence format	Pearson
Sequence type	aa
ClustalW version	1.82
JalView	
Output file	clustalw-20050511-23523458.output
Alignment file	clustalw-20050511-23523458.aln
Guide tree file	clustalw-20050511-23523458.dnd
Your input file	clustalw-20050511-23523458.input
<input type="button" value="SUBMIT ANOTHER JOB"/>	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

SeqA Name	Len(aa)	SeqB Name	Len(aa)	Score
=====	=====	=====	=====	=====
1 McHenry	454	2 Yurieva	426	97
=====	=====	=====	=====	=====

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Please check the out

Alignment

CLUSTAL W (1.82) multiple sequence alignment

```

McHenry      -SALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRRLAQAYLFSGPRGVGKTTTARLLAMAV 59
Yurieva      MSALYRRFRPLTFQEVG-QEHVKEPLLKAIRE-RLAQAYLFSGPRGVG-TTTARLLAMAV 57
*****

```

```

McHenry      GCQGEDPPCGVCPCQCQAVQGAHPDVVEIDAASNNNSVEDVRELRLRIHLAPLSAPRKVFI 119
Yurieva      GCQG-DPPCGVCPCQCQAVQR-AHPDVVDIDAASNNNS-EDVRELRLRIHLAPL-APRKVFI 113
          ****
          ****

McHenry      LDEAHMLSKSAFNALLKLTLEPPPHVLFVFATTEPERMPPTILSRTOHFRFRRLTEEEIA 179
Yurieva      LDEAHMLS-SAFNALLKLTLEPPP-VLFVFATTEPERMPP-ILSRTOHFRFRRLT-EEIA 169
          *****

McHenry      FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGLPP 239
Yurieva      FKLRRILEAVGR-AEEEALLLLARLADG-LRDAESLLERFLLLE-PLTRKEVERALGSPP 226
          *****

McHenry      REALAEIAASLARGKTAEALGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTPLP 299
Yurieva      -TGVAEIAASLARGKT-EALGLARRLYGEGYA-RSLVSGLLEVFREGL-AAFGLAGTPLP 282
          .:*****

McHenry      APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPPPEVGPKP 359
Yurieva      APPQ-LIAAMTALDEAMERLRR-SDALSLEVALLEA-RALAAEALPQPTGAPP-EVGPKP 338
          ****

McHenry      ESPPAPEPPRPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 419
Yurieva      ESPPTPEP-RPEEAPDLRERWRAF-EALRPTLRAFVREAR-EVREGQLCLAFPEDK-FHY 394
          ****:***

McHenry      RKASEQKARLLPLAQAHFGVEEVVLVLEGEKKSL 454
Yurieva      RKASEQKVRLLP-AQAHEFGVEEVVLVLEEKKKA-- 426
          *****:***

```

PLEASE NOTE: Showing colors on large alignments is slow.

Show Colors

View Alignment File

Guide Tree

Show as Phylogram Tree

Show Distances

View DND File

(McHenry:0.01291,Yurieva:0.01291);

Cladogram

McHe
Yuriev

Show as Phylogram Tree

Show Distances

View DND File

Right-click on the above tree to see display options.

Problems printing? Read [how to print a Phylogram or Cladogram](#).

[Return to site](#)

ClustalW Results

Results of search	
Number of sequences	2
Alignment score	2343
Sequence format	Pearson
Sequence type	aa
ClustalW version	1.82
JalView	
Output file	clustalw-20050511-23490919.output
Alignment file	clustalw-20050511-23490919.aln
Guide tree file	clustalw-20050511-23490919.dnd
Your input file	clustalw-20050511-23490919.input
<input type="button" value="SUBMIT ANOTHER JOB"/>	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

SeqA Name	Len(aa)	SeqB Name	Len(aa)	Score
1 McHenry	463	2 Yurieva	436	97

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Please check the out

Alignment

CLUSTAL W (1.82) multiple sequence alignment

```

McHenry      -SALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 59
Yurieva      MSALYRRFRPLTFQEVG-QEHVKEPLLKAIRE-RLAQAYLFSGPRGVG-TTTARLLAMAV 57
*****

```

McHenry	GCQGEDPPCGVCPHCQAVQGAHPDVVEIDAASNNVEDVRELREIRHLAPLSAPRKVFI	119
Yurieva	GCQG-DPPCGVCPHCQAVQR-AHPDVVDIDAASNN-EDVRELREIRHLAPL-APRKVFI	113

McHenry	LDEAHMLSsafnallktleppphvlfvfattefermpptilsrtqhfrrrlteeeia	179
Yurieva	LDEAHMLS-SAFNALLKTLEPPP-VLFVFATTEPERMPP-ILSRTQHFRFRRLT-EEIA	169

McHenry	FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGLPP	239
Yurieva	FKLRRILEAVGR-AEEEEALLLRLADG-LRDAESLLERFLLLE-PLTRKEVERALGSPP	226

McHenry	REALAEIAASLARGKTAEALGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTLP	299
Yurieva	-TGVAEIAASLARGKT-EALGLARRLYGEGY-RSLVSGLLEVFREGL-AAFGLAGTLP	282

McHenry	APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPPPEVGPKP	359
Yurieva	APPQ-LIAAMTALDEAMERLRR-SDALSLEVALLEA-RALAAEALPQPTGAPP-EVGPKP	338

McHenry	ESPPAPEPPRPEEAPDLRERWRFLAALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY	419
Yurieva	ESPPTPEP-RPEEAPDLRERWRFLAALRPTLRAFVREAR-EVREGQLCLAFPEDK-FHY	394

McHenry	RKASEQKARLLPLAQAHFGVEEVVLVLEGEKKSLEPKAPLGPTS	463
Yurieva	RKASEQKVRLLP-AQAHFGVEEVVLVLE-EKKSLEPKAPPGPTS	436

PLEASE NOTE: Showing colors on large alignments is slow.

Show Colors

View Alignment File

Guide Tree

Show as Phylogram Tree

Show Distances

View DND File

(McHenry:0.01032,Yurieva:0.01032);

Cladogram

McHe
Yurien

Show as Phylogram Tree

Show Distances

View DND File

Right-click on the above tree to see display options.

Problems printing? Read [how to print a Phylogram or Cladogram](#).

[Return to site](#)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.